Exercise training alters DNA-methylation patterns in genes related to muscle growth and differentiation in mice

Timo Kanzleiter1,2, *, Markus Jähnert1,2, Gunnar Schulze1,2, Joachim Selbig3, Nicole Hallahan1,2, Robert Wolfgang Schwenk1,2, Annette Schürmann1,2

1 Department of Experimental Diabetology, German Institute of Human Nutrition (DIFE), Potsdam-Rehbruecke, Germany
2 German Center for Diabetes Research (DZD)
3 Institute of Biochemistry and Biology and Institute of Computer Science / Bioinformatics University of Potsdam

* corresponding author.

Keywords:
DNA-methylation
Regular exercise training
Muscle development

Corresponding authors address:
Kanzleiter@dife.de
German Institute of Human Nutrition (DIFE)
Arthur-Scheunert-Allee 114-116
14558 Nuthetal, Germany
Abstract
The adaptive response of skeletal muscle to exercise training is tightly controlled and therefore requires transcriptional regulation. DNA-methylation is an epigenetic mechanism known to modulate gene expression but its contribution to exercise induced adaptations in skeletal muscle is not well-studied. Here we describe a genome-wide analysis of DNA-methylation in muscle of trained mice (n=3). In comparison to sedentary controls 2762 genes exhibited differentially methylated CpGs (p<0.05, meth diff >5%, coverage <10) in their putative promoter regions. Alignment with gene expression data (n=6) revealed 200 genes with a negative correlation between methylation and expression changes in response to exercise training. The majority of these genes were related to muscle growth and differentiation and a minor fraction involved in metabolic regulation. Among the candidates were genes that regulate the expression of myogenic regulatory factors (Plexin A2) as well as genes that participate in muscle hypertrophy (Igfbp4) and motor neuron innervation (Dok7). Interestingly, a transcription factor binding site enrichment study discovered significantly enriched occurrence of CpG methylation in the binding sites of the myogenic regulatory factors MyoD and myogenin. These findings suggest that DNA-methylation is involved in the regulation of muscle adaptation to regular exercise training.

Introduction
Skeletal muscle displays a remarkably high plasticity in its adaptive response to environmental stressors such as physical exercise. Importantly, one has to discriminate between acute and chronic adaptive responses. Very rapid responses such as activation of energy producing pathways (e.g. AMPK) or increases in insulin sensitivity are apparent immediately after a single bout of exercise, but only last for up to 48 hours (11, 37). In contrast chronic adaptations with long lasting effects such as muscle hypertrophy or increased oxidative capacity usually requires repeated exercise training. One striking difference is that these chronic responses require de novo gene expression and therefore underlie transcriptional regulation, whereas many acute effects do not require gene expression (37).

In recent years the contribution of epigenetic mechanisms has received increasing attention in the field of transcriptional regulation (20). The most intensely studied epigenetic modification so far is DNA-methylation. DNA-methylation usually occurs on cytosines followed by guanosines (CpG) that are most frequently found in
promoter regions and the first intron of genes. CpG methylation in these regulatory
regions is thought to lead to chromatin condensation and thereby block access of the
transcription machinery usually resulting in transcriptional silencing (6, 8).

This mechanism of gene regulation is important for such diverse processes as X
chromosome inactivation, cellular differentiation and growth of cancer cells (3, 9, 18).
However, the involvement of DNA methylation in the complex process of muscular
remodelling during periods of regular exercise training remains largely unexplored. A
recent study by Nitert and colleagues investigated genome wide DNA-methylation
patterns in first degree relatives with type 2 diabetes and control subjects without a
family history and found differences in DNA-methylation patterns that might be
involved in the onset of type 2 diabetes (32). Even more interesting they also
investigated the effect of a 6 months exercise intervention on the DNA-methylation
pattern and identified a number of genes involved in muscle metabolism that were
differently methylated after exercise. Surprisingly, even a single exercise bout seems
to have consequences on DNA-methylation patterns in skeletal muscle although they
seem to be of transient nature (2). Consequences of exercise on DNA-methylation
patterns appears not to be limited to skeletal muscle, but are also observable in
adipose tissue (38) and possibly other tissues as well, indicating that this a general
mechanism.

Getting further insight into epigenetic mechanisms regulating gene transcription
during exercise will help to further improve lifestyle interventions for, e.g., people
suffering from the metabolic syndrome. Therefore, the aim of our study was to
investigate the role of DNA-methylation in adaptation of skeletal muscle to regular
exercise training.

**Material and Methods**

**Animal studies**

Animals were kept in a temperature-controlled room (22±1°C) on a 12 h light/dark
cycle with free access to food and water. All animal studies were conducted in
accordance with the FELASA guidelines for the care and use of laboratory animals
(14), and all experiments were approved by the ethics committee of the State Agency
of Environment, Health and Consumer Protection (State of Brandenburg, Germany)
under permit number V3-2347-06-2011.
**Exercise training**

Nine week old male C57BL/6J mice (Charles River, Germany) were randomly assigned to the exercise (n=24; EX) or sedentary (n=24; SED) group. Endurance training was performed on a motorized treadmill (Exer6; Columbus Instruments; Columbus, USA) on 5 days followed by 2 days of rest for 4 weeks. The treadmill was equipped with an electrical shock grid to encourage mice to complete the training sessions. To ensure a continuous training challenge a progressive training protocol, which increased its intensity over the four week training period from 30 to 50 minutes per day, was used (Table 1). Mice were removed from the study if they failed to complete two training sessions. The criterion for failure was that the mouse remained on the shock grid for 5 consecutive seconds and would not resume exercise after manual prodding. The shock grid was then turned off and the mouse was removed from the treadmill. This was the case for 4 out of 24 mice that underwent the training regime and these four mice were therefore removed from the study. Three hours after completion of the last training session mice were sacrificed and quadriceps muscles were rapidly dissected and frozen in liquid nitrogen. In order to avoid sampling bias due to different physiological properties of the quadriceps muscle whole muscles were powdered and aliquots were used for subsequent measurement.

**Body composition and weight measurements**

Body composition was analyzed weekly before the exercise training started by nuclear magnetic resonance (NMR) (Minispec LF50; Bruker Biospin Corporation, Billerica, MA, USA). Body weight measurements were started the week before the first exercise training and then weekly recorded on the same day as body composition was determined.

**Citrate synthase assay**

Citrate synthase activity was measured in the muscle of the mice as a surrogate measurement for changes in oxidative capacity in response to exercise training. Assays were performed as described previously (21). Briefly, frozen quadriceps muscle samples were homogenized in 19 volumes of Tris-EDTA buffer (pH 7.4), and cleared supernatant after centrifugation was used for a spectrophotometric assay. The conversion rate of acetyl coenzyme A (acetyl-CoA) and oxaloacetate to citrate and CoA-SH by citrate synthase is proportional to the coupled reaction of CoA-SH and 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB; Sigma-Aldrich) to 2-nitro-5-
thiobenzoate (NTB), which was measured at 412 nm. The activity of citrate synthase in muscle is expressed per mg muscle wet weight.

**Reduced representation bisulfite sequencing (RRBS)**
Genomic DNA isolated from muscle samples (n=3/group) was used to construct RRBS-libraries for sequencing according to published protocols (13, 26). In brief, genomic DNA was digested by the restriction enzyme MspI to generate short fragments that contain CpG dinucleotides at the ends. After end-repair, A-tailing and ligation to methylated Illumina adapters, the CpG-enriched DNA fragments were size selected, subjected to bisulfite conversion, amplified and end sequenced on an Illumina HiSeq2000. Library construction and sequencing was performed by Zymo Research Corporation (USA).

Reads were quality controlled with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimmed to remove adapter sequences with Trim Galore (www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Next, processed reads were mapped to the *in silico* bisulfite converted reference genome (mm9; Build37) using Bismark software (22, 23). CpG methylation calls from Bismark were filtered with a custom PERL-script that selects only CpGs within a region 2kb upstream and 1kb downstream of UCSC annotated transcriptional start sites (mm9; Build37).

The methylation ratio (Mr) is defined as the number of reads overlapping a particular CpG comprising a C or a T at the first position of the CpG. If x is the number of C’s and y the number of T’s the formula $Mr = x/(x+y)$ then gives the methylation ratio for each CpG. Only CpGs with a sequencing depth of $>10x$ and a methylation difference $>5\%$ in all samples were considered for further analysis. The methylation data has been deposited together with the expression profiling as a super series on GEO (GSE54278).

**Transcriptional profiling**
RNA was isolated from quadriceps muscle samples (n=6/group) according to the manufacturer’s instructions using TRI-reagent (Sigma Aldrich). RNA integrity was measured on LabChips in an Agilent Bioanalyzer (Agilent) and concentration was determined using a Nanodrop spectrophotometer (Thermo Scientific). Only samples with a RNA integrity number higher than 9 were used for array hybridization. Transcriptional profiling of muscle samples were performed on Agilent 4x44k whole genome microarrays by Source Biosciences (Berlin).
Validation of differential gene expression by quantitative real time PCR

For validation of gene expression in candidates identified in our transcriptome analysis RT-PCR was performed in muscle samples from exercised (EX; n=16) and sedentary control mice (SED; n=20). RNA concentration was determined spectrophotometrically using a Nanodrop spectrophotometer (Thermo Scientific). For conversion into cDNA random primers and the cDNA synthesis kit from Qiagen were used. Quantitative real time PCR was performed on Roche Lightcycler 480 using gene specific primers and corresponding UPL Probes (Roche) (Table S3). The eukaryotic translation elongation factor 2 (Eef2) was used as endogenous control.

Transcription factor binding site (TFBS) enrichment

Narrow peak regions of TFBS for all 12 available transcription factors in the C2C12 cell line were downloaded from the ENCODE Transcription Factor Binding Track. CpGs that mapped within an annotated TFBS were considered a ‘hit’ for the respective transcription factor. Significance of enrichment was calculated with Fischer’s exact test using R (43).

Statistical tests

All statistical tests were performed in SPSS. For parametric tests data was first tested for normality of distribution using Shapiro-Wilk test and for homogeneity of variances with Levene test.

Significance of difference for body composition and –weight (Fig 1a-c) were calculated with one-way ANOVA and pairwise comparison was performed with posthoc Sidak test. Difference in citrate synthase activity (Fig1 d) was tested with students t-test.

Data from transcriptional profiling and methylation analysis were tested with student’s t-test. Correction for multiple testing with common methods such as Bonferroni or FDR-based methods such as Benjamini-Hochberg correction tend to be overly conservative and reduce false positives on the expense of false negatives (15, 34, 39). Therefore, we decided against correction for multiple testing with indicated methods. Instead we assumed a causal relation between DNA-methylation in promoter regions and changes in expression of the respective gene. We applied a strategy of connecting our two datasets by generating overlaps between expression and methylation data to create a “biological significance”. The idea behind this is that an observation made purely by chance (type-1-error) is much less likely to occur in
the overlap of the two datasets. Therefore, this strategy stratifies for false positive results, as would correction for multiple testing without the very high stringency of mathematical corrections as discussed elsewhere (15).

For hierarchical clustering the values from RRBS analysis and transcriptional profiling have been normalized (log transformation and subtraction of mean values for respective gene). Clustering and generation of heatmap was performed with ‘R’.

The calculation of enrichment of transcription factor binding sites in differentially methylated regions was performed using Fischer’s exact test.

**Results**

**Training regime**

Exercise training resulted in a significantly reduced body fat content from week 2 on (Fig 1A), whereas lean mass was slightly lower in the first two weeks of the study (Fig 1B). Body weight tended to be lower in trained mice (n=20) as compared to sedentary mice (n=16; Fig 1C) most likely due to reduced body fat. As an additional readout of training success we measured the activity of the mitochondrial enzyme citrate synthase in quadriceps muscle homogenates (n=11), which was significantly increased after 4 weeks of exercise training (Fig 1D).

**Impact of exercise training on DNA methylation**

Reduced representation bisulfite sequencing was used to assess changes in the DNA-methylation levels in response to regular exercise. Quadriceps muscle samples taken three hours after completion of the last training session were used for the construction of RRBS libraries (n=3/group), as many genes expressed in response to adaptation to regular exercise training in the muscle still show differential expression after three hours. Only sequencing data from CpGs that were located within the promoter region, defined as 2kb upstream and 1kb downstream of the transcriptional start site (Fig 2A), and had sequencing depth of at least 10x for all samples were included in the analysis. Using these criteria our sequencing data from RRBS fragments covered about 55% of all promoter regions in the genome (Fig 2B). In total we detected 3692 differentially methylated CpGs with at least 5% difference in methylation levels (p<0.05, meth-diff >5%, coverage <10) that were distributed among 2762 promoter regions (Fig 2A).
Comparison of transcriptional and DNA methylation changes

Transcriptional changes in response to exercise were assessed in muscle samples from an increased group of animals (n=6/group) comprising the same animals used for methylation analysis (n=3) and additional three animals that were randomly picked from the trained or sedentary group. The analysis revealed 3020 differentially expressed genes (p<0.05) between sedentary and exercised animals. Comparison of expression and methylation data revealed 361 genes (479 CpGs) in which differential DNA methylation were also associated with differential gene expression.

Hierarchical clustering showed little variation in expression and methylation pattern between animals of the same treatment group (exercise or sedentary) and thereby confirmed results of the principal component analysis. The heatmap also revealed at least four distinct clusters (Fig 3B). Specifically, cluster 1 and cluster 4 represent genes that show decreased or increased expression and DNA methylation in response to exercise, respectively. Whereas cluster 2 and 3 consist of genes with increased expression and decreased DNA methylation or vice versa. The two clusters with negative correlation of expression and methylation levels (cluster 2 and 3; Fig 3B) include 200 genes. Of those 66 genes were hypomethylated and had increased expression and 134 genes were hypermethylated in response to exercise and had a decreased expression (Table S1).

Biological pathways affected by regular exercise

Next we performed KEGG pathway analysis using WebGestalt (http://bioinfo.vanderbilt.edu/webgestalt/) to identify biological pathways that were affected by exercise (Table S2). Analysis of our methylation data revealed significant enrichment in pathways involved in growth and differentiation (Fig 4A, bold font), such as the MAPK signaling pathway or TGF-beta signaling pathway. Metabolic pathways were also significantly enriched (Fig 4A, italic font), such as the KEGG category Metabolic pathways, insulin signaling pathway or oxidative phosphorylation. Pathway analysis of gene expression data showed a similar picture with the top three pathways from methylation analysis (Metabolic pathways, MAPK signaling pathway and Wnt signaling pathway) also being significantly enriched (Fig 4B). Analysis of the 200 genes that showed a negative correlation between DNA methylation and gene expression revealed mainly pathways connected to growth and differentiation such as Axon guidance and MAPK signaling pathway, but also two metabolic pathways.
(Metabolic pathways and insulin signaling pathway) already found in the methylation only data (Fig 4C).

**Differentially methylated and expressed genes**

Among the genes with significant changes in promoter methylation and gene expression was a regulator of myogenin Plexin A2 (Plxna2) (Fig 5A). The same changes in methylation and expression pattern was also observed in the growth and differentiation related genes insulin-like growth factor binding protein 4 (Igfbp4), docking protein 7 (Dok7) and CDP-diacylglycerol synthase (Cds2) (Fig 5A).

Negative correlation between methylation and expression was also observed in a number of genes with functions unrelated to growth and differentiation. Calcium binding protein 39 (Cab39) is involved in the activation of the AMPK pathway (16), START domain containing 10 (Stard10) influences phospholipid content of mitochondrial membranes and thereby oxidative phosphorylation (44), Slc20a1 is a sodium phosphate co-transporter and glutamine fructose-6-phosphate transaminase 2 (Gfp2) is involved in the regulation of glycogen storage (Fig 6A) (30).

**Enrichment of transcription factor binding sites (TFBS)**

DNA-methylation influences gene transcription by condensing chromatin and thereby prevents access of transcription factors. Therefore, we investigated whether the binding sites for any of the 12 transcription factors annotated in the skeletal muscle cell line C2C12 in the recently published ENCODE data set (5, 12, 31) is enriched for differentially methylated CpGs from our study. TFBS for factors MyoD, myogenin, Max, TCF3, CTCF and E2F4 were significantly enriched (Table 2). Interestingly, the highest significance was observed for the myogenic regulatory factors MyoD and myogenin, which are main regulators of muscle growth and differentiation.

**Discussion**

Physical inactivity contributes to lifestyle-related diseases including obesity, type 2 diabetes, hypertension, heart disease and age-related muscle wasting. Although the benefits of regular exercise are long known, the underlying regulatory mechanisms are still not completely understood. Recently, first evidence was published that epigenetic mechanisms might be involved in muscle adaptation to regular exercise training in men with a family history of type 2 diabetes (32).

In our study we used inbred C57BL/6J mice and endurance trained them for four weeks on a motorized treadmill. This experimental approach minimizes contributions
from differences in genetic background and allows for a tightly controlled training intensity. Trained mice displayed a decreased fat mass and increased oxidative capacity as to be expected in response to regular exercise (35). Different muscle groups respond with different intensities to exercise training. However, for the analysis of genome wide DNA-methylation we used the quadriceps muscle, because (a) it is commonly used in rodent exercise studies, (b) it is sufficient in size to perform all measurements in the same muscle and (c) the available studies on effects of exercise on DNA-methylation in humans used the vastus lateralis muscle, which is part of the quadriceps muscle. We performed reduced representation bisulfite sequencing, as it was shown to have a higher likelihood of high coverage in promoter regions as compared to MeDIP-seq or MethylCap-seq (7). In our study 55% of all promoter regions were covered, which is comparable to previously published studies (7). We detected 3692 differentially methylated CpGs that were distributed among 2762 promoter regions. Our main interest in this study was to identify consequences of DNA-methylation on gene expression in response to regular exercise training. Direct comparison of methylation and expression data resulted in a list of 361 differentially methylated and expressed genes. Since DNA-methylation constitutes an epigenetic mark usually correlated with transcriptionally silent condensed chromatin we filtered our gene list accordingly and identified 200 genes with a negative correlation between methylation and expression levels. However, the authors have to point out that this study must be seen as an exploratory approach to test the involvement of DNA-methylation in post-exercise adaptation. Due to the relatively low sample number (n=3) used for the genome wide analysis of DNA-methylation by RRBS standard procedures used for correction of multiple testing errors could not be employed. Conclusions based on the methylation data only are therefore subject to a certain degree of statistical uncertainty. In an effort to at least partially overcome this fact we correlated the methylation data with genome wide expression data and single gene expression validation by RT-PCR with higher sample numbers to add a biological significance as discussed in more detail in the method section. KEGG pathway analysis of differentially methylated genes revealed that within the five most significantly enriched pathways four were related to growth and differentiation and one to metabolic regulation (Fig 4A). The analysis of genes that showed negative correlation between methylation and expression revealed a similar
result with three of the top five enriched pathways being related to growth and differentiation (Fig 4C). The majority of published gene expression studies found that growth related pathways only made up a small part of the affected biological functions and the majority of genes had a function in metabolic adaptation (reviewed here (10)).

The dominance of growth and differentiation associated pathways prompted us to have a closer look at differentially methylated genes with a known function in muscle growth and differentiation.

Noteworthy, Plexin A2 is part of the Semaphorin A receptor and able to induce myogenin expression in response to Semaphorin A binding during muscle development (41).

Insulin like growth factor binding protein 4 (Igfbp4) was identified as a muscle growth related gene that was induced by hypertrophic stimuli, such as exercise, and decreased under atrophic conditions (1). Consistent with its putative function in muscle hypertrophy we found Igfbp4 to be hypomethylated and induced after exercise training.

Muscle growth does not only depend on growth and differentiation of muscle cells but also requires processes such as motor neuron innervation and supply by blood vessels. Interestingly, Dok7, an activator of muscle-specific receptor kinase, is involved in the assembly of the postsynaptic neuromuscular junction at the endplate area of myotubes (33). The increased expression and decreased methylation in response to exercise might indicate that Dok7 is involved in muscle hypertrophy or repair. In addition, CDP-diacylglycerol synthetase 2 (Cds2) regulates vascular endothelial growth factor signaling and thereby affects angiogenesis. Cds2 also showed significantly reduced methylation and increased expression consistent with a possibly increased rate of angiogenesis in response to exercise.

If muscle growth and differentiation in response to regular exercise training is indeed influenced by DNA-methylation (Fig. 5B) then one would expect the major pathways to be affected. The family of myogenic regulatory factors contains the four main factors involved in muscle growth and differentiation (MyoD, myogenin, Myf5 and MRF4). Within this gene family MyoD has the function of a master regulator shown by its ability to drive conversion of a variety of cell lineages into myogenic cells (25).
murine muscle cells we found that the occurrence of differentially methylated CpGs near the binding sites for MyoD and myogenin displayed the highest significance. This could indicate that many downstream targets of MyoD and Myogenin are affected through changes in DNA-methylation by altering accessibility of their respective binding sites.

Noteworthy, treatment of myoblasts with the DNA-methylation inhibitor 5-azacytidine was shown to promote myogenesis (19), which is consistent with our observation that the majority of growth and differentiation related genes in our list show decreased methylation.

Interestingly, we could not detect changes in DNA-methylation of contractile genes such as myosin heavy chains, troponin etc. Therefore, we can only conclude that at least under the conditions used in our study these crucial components of muscle adaption to exercise are not regulated by changes in DNA-methylation.

Although the majority of the 200 genes with a negative correlation between methylation and expression were related to growth we also found several genes with a function in muscle metabolism (Fig. 6B). One particularly interesting candidate we identified is Cab39, which activates LKB1 the major upstream kinase of the AMP-activated protein kinase (AMPK) (16, 40). AMPK is a master metabolic regulator responsible for increased glucose uptake (17), increased fatty acid oxidation (27) and induction of mitochondrial biogenesis (4) in response to an energy challenges such as exercise. Another regulator of lipid metabolism among our candidates is Stard10, which modulates phospholipid content of mitochondrial membranes and thereby influences mitochondrial function (44). Cellular energy metabolism depends on high-energy phosphate bonds such as in ATP and phosphocreatine for short-term energy storage or inorganic phosphate as substrates for tricarboxylic acid cycle and mitochondrial FoF1 ATPase. Consistent with these requirements we found the sodium phosphate co-transporter Slc20a1 had a higher expression and lower methylation in response to exercise. Another mechanism of energy storage in muscle is the condensation of glucose molecules into large multibranched glycogen molecules. Gfpt2 is the rate-limiting enzyme of the hexosamine biosynthesis pathway and increased flux through this pathway was postulated as a mechanism to limit excessive post-exercise glycogen deposition (30).

The early studies on epigenetic regulation of gene expression regarded histone modifications as short-term and DNA-methylation as long-term epigenetic marks.
(reviewed here (36)). And indeed there is a large body of literature showing that in many cases DNA-methylation is not only maintained throughout cell division but is even heritable. Although many of the beneficial effects of exercise are short lived, there are a few longitudinal studies indicating that even after many decades of training cessation former athletes still have a lower risk for glucose intolerance or type 2 diabetes (24, 42). Importantly, this was independent of current physical activity levels. These long-term consequences of exercise might be at least in part due to changes in the DNA-methylation pattern.

Taken together our results suggest that genes involved in many different aspects of muscle growth and metabolic adaptation are subject to regulation by DNA-methylation in response to regular exercise. This might lead to the development of novel therapies that could mimic some effects of exercise to treat disturbances of muscle growth such as muscular dystrophy or provide longer lasting beneficial metabolic effects. However, due to the exploratory nature of this study the identified candidates should be subject to confirmation in future studies.
References


trez&rendertype=abstract [10 Sep. 2014].

trez&rendertype=abstract [28 Nov. 2013].


t=Citation&list_uids=18380005.


Figure legends

**Figure 1: Effects of exercise training on body composition and on muscle citrate synthase activity**

Four weeks of progressive exercise training resulted in significantly reduced (A) fat mass, whereas (B) lean mass and (C) body mass were only slightly reduced (*P<0.05; n=16-20; One-way ANOVA, post hoc Kruskal Wallis test). (D) Activity of the mitochondrial enzyme citrate synthase in muscle (#P<0.05; n=11 students t-test). (E) Number of mice included in different assays.

**Figure 2: DNA-methylation analysis**

(A) The genomic region 2kb upstream and 1kb downstream of transcriptional start sites (TSS) was defined as putative promoter region. Differentially methylated CpGs (CG) identified by RRBS analysis that were located within these regions were considered for further analysis if they had >5% difference in methylation and sequence depth of >10x. (B) Coverage and sequence depth of putative promoter-regions from all 24361 annotated genes. Numbers within pie chart indicate portion of all promoter-regions in percent and numbers next to chart indicate sequence depth. About 55% of putative promoter regions were present in our RRBS analysis with a sequence depth higher than 10x.

**Figure 3: Comparison of changes in expression and methylation**

(A) Venn diagram of differentially methylated and expressed genes. (B) Hierarchical clustering of the 361 genes present in both data sets revealed two clusters (cluster 2&3) with a negative correlation between expression and methylation containing 200 genes. Blue color symbolizes a decrease and red color an increase in DNA-methylation or gene expression.

**Figure 4: Biological pathways affected by regular exercise training**

KEGG pathway analysis was performed on (A) differentially methylated and (B) differentially expressed genes and (C) genes with a negative correlation between expression and methylation. Pathways in bold font are related to growth and differentiation and metabolic pathways are shown in italic font. Numbers in brackets...
indicate number of genes in the reference category, numbers in bars represent the ratio of enrichment and adjusted P (adjP) indicates the significance of enrichment.

**Figure 5: Effects of regular exercise on methylation and expression of selected genes related to muscle growth**

(A) Percent methylation (RRBS; n=3; *p<0.05; students t-test) and mRNA abundance (RT-PCR; n=16-20; *p<0.05; students t-test) in quadriceps muscle samples of genes related to muscle growth and differentiation. SED: sedentary, EX: exercised (B) Biological function and possible connection of selected candidates.

**Figure 6: Effects of regular exercise on methylation and expression of selected metabolic genes**

(A) Percent methylation (RRBS; n=3; *p<0.05; students t-test) and mRNA abundance (RT-PCR; n=16-20; *p<0.05; students t-test) in quadriceps muscle samples of genes with a known metabolic function. SED: sedentary, EX: exercised (B) Biological function and possible connection of selected candidates.
Table 1: Progressive training protocol (5 days/week)

<table>
<thead>
<tr>
<th></th>
<th>Week1</th>
<th>Week2</th>
<th>Week3</th>
<th>Week4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incline</td>
<td>0°</td>
<td>0-5°</td>
<td>5-10°</td>
<td>10°</td>
</tr>
<tr>
<td>Daily duration</td>
<td>30-40min</td>
<td>40min</td>
<td>40min</td>
<td>40-50min</td>
</tr>
<tr>
<td>Max speed</td>
<td>15-18m/min</td>
<td>18m/min</td>
<td>18m/min</td>
<td>18m/min</td>
</tr>
</tbody>
</table>
**Table 2:**

**Differentially methylated CpGs in transcription factor binding sites**

<table>
<thead>
<tr>
<th>Transcription factor</th>
<th>Expected (%)</th>
<th>Observed (%)</th>
<th>Significance of enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td>MyoD</td>
<td>19.49</td>
<td>37.28</td>
<td>p=6.2 x 10^{-299}</td>
</tr>
<tr>
<td>Myogenin</td>
<td>39.94</td>
<td>78.32</td>
<td>p=1.0 x 10^{-300}</td>
</tr>
<tr>
<td>Max</td>
<td>16.83</td>
<td>23.16</td>
<td>p=6.12 x 10^{-46}</td>
</tr>
<tr>
<td>TCF3</td>
<td>6.85</td>
<td>12.95</td>
<td>p=2.86 x 10^{-85}</td>
</tr>
<tr>
<td>CTCF</td>
<td>15.45</td>
<td>21.42</td>
<td>p=8.51 x 10^{-44}</td>
</tr>
<tr>
<td>E2F4</td>
<td>14.87</td>
<td>17.04</td>
<td>p=2.09 x 10^{-7}</td>
</tr>
</tbody>
</table>
Figure 1
**Figure 2**

**A**

Putative promoter

2kb  |  1kb

CG    |  CG    |  CG

TSS

3692 CpGs
(p<0.05; meth.diff >5%; coverage >10x) in

2762 promoter regions

**B**

Coverage of promoter regions
(from 24361 annotated genes)

- >=100x: 43%
- <100x: 45%
- <50x: 5%
- <25x:
Figure 4

A DNA-methylation
(2762 genes; p<0.05; meth.diff >5%)

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Expected</th>
<th>Observed</th>
<th>adjP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metabolic pathways (1184)</td>
<td></td>
<td>2.58</td>
<td>adjP = 9.02 x 10^-23</td>
</tr>
<tr>
<td>MAPK signaling pathway (268)</td>
<td></td>
<td>4.84</td>
<td>adjP = 9.02 x 10^-23</td>
</tr>
<tr>
<td>Wnt signaling pathway (154)</td>
<td></td>
<td>5.62</td>
<td>adjP = 9.24 x 10^-18</td>
</tr>
<tr>
<td>TGF-beta signaling pathway (85)</td>
<td></td>
<td>6.36</td>
<td>adjP = 6.0 x 10^-13</td>
</tr>
<tr>
<td>Cell cycle (127)</td>
<td></td>
<td>5.87</td>
<td>adjP = 6.96 x 10^-12</td>
</tr>
<tr>
<td>Insulin signaling pathway (137)</td>
<td></td>
<td>4.74</td>
<td>adjP = 7.3 x 10^-12</td>
</tr>
<tr>
<td>Calcium signaling pathway (178)</td>
<td></td>
<td>3.40</td>
<td>adjP = 5.91 x 10^-8</td>
</tr>
<tr>
<td>mTOR signaling pathway (53)</td>
<td></td>
<td>5.71</td>
<td>adjP = 2.64 x 10^-7</td>
</tr>
<tr>
<td>Type 2 diabetes mellitus (49)</td>
<td></td>
<td>4.86</td>
<td>adjP = 2.92 x 10^-5</td>
</tr>
<tr>
<td>Oxidative phosphorylation (147)</td>
<td></td>
<td>2.35</td>
<td>adjP = 0.0023</td>
</tr>
<tr>
<td>Citrate cycle (TCA cycle) (31)</td>
<td></td>
<td>4.19</td>
<td>adjP = 0.0044</td>
</tr>
</tbody>
</table>

B Gene expression
(3020 genes; p<0.05)

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Expected</th>
<th>Observed</th>
<th>adjP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathways in cancer (325)</td>
<td></td>
<td>6.06</td>
<td>adjP = 9.43 x 10^-14</td>
</tr>
<tr>
<td>Wnt signaling pathway (154)</td>
<td></td>
<td>6.00</td>
<td>adjP = 1.08 x 10^-6</td>
</tr>
<tr>
<td>mRNA surveillance pathway (93)</td>
<td></td>
<td>7.94</td>
<td>adjP = 1.08 x 10^-6</td>
</tr>
<tr>
<td>Neurotrophin signaling pathway (131)</td>
<td></td>
<td>6.58</td>
<td>adjP = 1.08 x 10^-6</td>
</tr>
<tr>
<td>Metabolic pathways (1184)</td>
<td></td>
<td>6.58</td>
<td>adjP = 1.54 x 10^-6</td>
</tr>
<tr>
<td>MAPK signaling pathway (268)</td>
<td></td>
<td>4.38</td>
<td>adjP = 1.96 x 10^-6</td>
</tr>
<tr>
<td>Chemokine signaling pathway (185)</td>
<td></td>
<td>4.99</td>
<td>adjP = 5.06 x 10^-6</td>
</tr>
<tr>
<td>Apoptosis (86)</td>
<td></td>
<td>6.44</td>
<td>adjP = 9.76 x 10^-5</td>
</tr>
<tr>
<td>Endocytosis (220)</td>
<td></td>
<td>3.92</td>
<td>adjP = 0.0001</td>
</tr>
<tr>
<td>Notch signaling pathway (50)</td>
<td></td>
<td>8.62</td>
<td>adjP = 0.0001</td>
</tr>
<tr>
<td>Purine metabolism (168)</td>
<td></td>
<td>4.40</td>
<td>adjP = 0.0001</td>
</tr>
</tbody>
</table>

C DNA-methylation & gene expression
(200 genes; p<0.05; negative correlation)

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Expected</th>
<th>Observed</th>
<th>adjP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axon guidance (131)</td>
<td></td>
<td>11.15</td>
<td>adjP = 0.0006</td>
</tr>
<tr>
<td>Dorso-ventral axis formation (22)</td>
<td></td>
<td>39.85</td>
<td>adjP = 0.0006</td>
</tr>
<tr>
<td>Neurotrophin signaling pathway (131)</td>
<td></td>
<td>11.15</td>
<td>adjP = 0.0006</td>
</tr>
<tr>
<td>B cell receptor signaling pathway (76)</td>
<td></td>
<td>15.38</td>
<td>adjP = 0.0006</td>
</tr>
<tr>
<td>Hepatitis C (137)</td>
<td></td>
<td>10.67</td>
<td>adjP = 0.0006</td>
</tr>
<tr>
<td>GnRH signaling pathway (99)</td>
<td></td>
<td>11.81</td>
<td>adjP = 0.0006</td>
</tr>
<tr>
<td>Insulin signaling pathway (137)</td>
<td></td>
<td>8.53</td>
<td>adjP = 0.0066</td>
</tr>
<tr>
<td>Fc epsilon RI signaling pathway (80)</td>
<td></td>
<td>10.98</td>
<td>adjP = 0.0098</td>
</tr>
<tr>
<td>Metabolic pathways (1184)</td>
<td></td>
<td>2.71</td>
<td>adjP = 0.0098</td>
</tr>
<tr>
<td>MAPK signaling pathway (268)</td>
<td></td>
<td>5.45</td>
<td>adjP = 0.0098</td>
</tr>
<tr>
<td>VEGF signaling pathway (76)</td>
<td></td>
<td>11.54</td>
<td>adjP = 0.0098</td>
</tr>
</tbody>
</table>
Figure 5

A

Plxna2

Igfbp4

Dok7

Cds2

B

Igfbp4  Plxna2  Dok7  Cds2

MyoD  myogenin

TFBS differentially methylated

Muscle growth and differentiation  Muscle innervation  Angiogenesis
Figure 6

A

- **Cab39**
  - 5meCitolysines (%)
  - Cab39 mRNA (AU)

- **Stard10**
  - 5meCitolysines (%)
  - Stard10 mRNA (AU)

- **Gfpt2**
  - 5meCitolysines (%)
  - Gfpt2 mRNA (AU)

- **Slc20a1**
  - 5meCitolysines (%)
  - Slc20a1 mRNA (AU)

B

- **Gfpt2**
  - Glycogen deposition
- **Cab39**
  - AMPK activation
- **Stard10**
  - Membrane phospholipids
- **Slc20a1**
  - Phosphate import

- Glucose metabolism
- Mitochondrial lipid metabolism